1.0

1.1

1.25   1.4   1.6

28   2.5

2.2

2.0

1.8

LEVEL

INSTANTANEOUS CONTROL OF BROWNIAN MOTION

by

J. Michael Harrison
Graduate School of Business
Stanford University

and

Michael I. Taksar
Department of Operations Research
Stanford University

DTIC
ELECTE
JAN 20 1982
B

DEPARTMENT OF OPERATIONS RESEARCH
AND
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Instantaneous Control of Brownian Motion

J. Michael Harrison
Michael I. Taksar

Stanford University

Abstract

A controller continuously monitors a storage system, such as an inventory or bank account, whose content $Z = \{Z_t, t \geq 0\}$ fluctuates as a (μ, σ^2) Brownian motion in the absence of control. Holding costs are incurred continuously at rate $h(Z_t)$. At any time, the controller may instantaneously increase the content of the system, incurring a proporitional cost of r times the size of the increase, or decrease the content at a cost of ℓ times the size of the decrease. We consider the case where h is convex on a finite interval $[\alpha, \beta]$ and $h = \infty$ outside this interval. The objective is to minimize the expected discounted sum of holding costs and control costs over an infinite planning horizon.

It is shown that there exists an optimal control limit policy, characterized by two parameters a and b ($\alpha \leq a < b \leq \beta$). Roughly speaking, this policy exerts the minimum amounts of control sufficient to keep $Z_t \in [a,b]$ for all $t \geq 0$. Put another way, the optimal control limit policy imposes on Z a lower reflecting barrier at a and an upper reflecting barrier at b. We do not give a full-blown algorithm for construction of the optimal control limits, but a computational scheme could easily be developed from our constructive proof of existence.

Instantaneous Control of Brownian Motion

J. Michael Harrison
Michael I. Taksar

Stanford University

1. <u>Introduction</u>

Consider a controller who continuously monitors the content of a storage system, such as an inventory or bank account. In the absence of any control, the content process $Z = \{Z_t, t \geq 0\}$ fluctuates as a Brownian Motion with drift μ and variance σ^2, and <u>holding costs</u> are continuously incurred at rate $h(Z_t)$. In order to avoid excessive holding costs, the controller may at any time increase the content of the system by any amount desired, incurring a proportional cost of r times the size of the increase. Similarly, he may decrease the content by any amount desired, incurring a proportional cost of ℓ times the size of the decrease. Hereafter we use the term <u>pushing right</u> to mean increasing the content of the system, and <u>pushing left</u> to mean decreasing the content. Thus r and ℓ are the proportional control costs associated with pushing right and pushing left respectively.

The controller's objective is to find a policy that minimizes the expected discounted sum of holding costs and control costs over an infinite planning horizon, where future costs are continuously

discounted at interest rate $\gamma > 0$. To formulate this problem in precise mathematical terms, we begin with a (μ, σ^2) Brownian motion $X = \{X_t, t \geq 0\}$, denoting by P_x the distribution on the path space of X corresponding to initial state x. A policy is defined as a pair of nonnegative processes $R = \{R_t, t \geq 0\}$ and $L = \{L_t, t \geq 0\}$ that are non-decreasing and non-anticipating with respect to X. Interpret R_t and L_t as the <u>cumulative</u> amounts of rightward movement and leftward movement, respectively, effected by the controller over the time interval $[0,t]$. The content process under policy (R,L) is

$$Z_t = X_t + R_t - L_t , \qquad t \geq 0 ,$$

and we define the associated cost function

$$k(x) = E_x \left[\int_0^\infty e^{-\gamma t} h(Z_t) dt + r \int_0^\infty e^{-\gamma t} dR_t + \ell \int_0^\infty e^{-\gamma t} dL_t \right] ,$$

with the Riemann-Stieltjes integrals on the right defined to include the control costs rR_0 and ℓL_0 incurred at $t = 0$ (see §3). Our objective is to find a policy which minimizes $k(x)$ for every starting state x.

An essential feature of this problem is that the controller can <u>instantaneously</u> change the content (or state) of the storage system. Thus, it is possible to further impose <u>state constraints</u> on the controller's actions, which may be formally expressed by setting $h(x) = \infty$ for some states x. In the same way, one of the controller's options may be eliminated by setting $r = \infty$ or $\ell = \infty$. The special case where $r < \infty$, $\ell < \infty$ and

2

$$(1.1) \qquad h(x) = \begin{cases} x & \text{if } x \geq 0 \\ \\ \infty & \text{if } x < 0 \end{cases}$$

was studied by Harrison and Taylor [6]. Defining b as the unique solution of a certain transcendental equation, they proved the optimality of a <u>control limit policy</u> (R^*, L^*) with lower limit zero and upper limit b. The policy (R^*, L^*) and its associated content process $Z^* = X + R^* - L^*$ may be described as follows. If $X_0 < 0$, one takes $R_0^* = -X_0$, so that $Z_0^* = 0$. If $X_0 > b$, one takes $L_0^* = X_0 - b$, so that $Z_0^* = b$. After time zero, one increases R^* and L^* in the minimal amounts sufficient to achieve $0 \leq Z_t^* \leq b$ for all $t \geq 0$. Under this plan, Z^* is a (μ, σ^2) Brownian motion with a lower reflecting barrier at zero and an upper reflecting barrier at b, $R^* - R_0^*$ is the <u>local time</u> of Z^* at zero, and $L^* - L_0^*$ is the local time of Z^* at b. In particular, although R^* and L^* may have jumps at $t = 0$, they are continuous but <u>singular</u> thereafter. The last phrase means that the set of time points at which R^* or L^* increases has zero Lebesgue meansure (almost surely). This singularity expresses a bang-bang property of the optimal policy with *instantaneous control.*

In this paper we consider the instantaneous control problem with $r < \infty$, $\ell < \infty$ and a holding cost function of the form

3

$$h(x) = \begin{cases} \infty \, , & x < \alpha \\ \text{a general convex function} \, , & \alpha \leq x \leq \beta \\ \infty \, , & x > \beta \end{cases}$$

where $-\infty < \alpha < \beta < \infty$. It will be shown that there exists an optimal control limit policy with lower limit a and upper limit b, where $\alpha \leq a < b \leq \beta$. We do not present a full-blown algorithm for computation of the optimal control limits a and b, but a computational scheme could easily be developed from our constructive proof of existence. Our treatment generalizes the result by Harrison and Taylor [6] described earlier, and the methods used here are also more elegant and more general in their applicability. This improvement in methodology and presentation has itself been a major goal in our study, although the extension to general convex holding costs is potentially important for applications.

It will ultimately be found that the minimal cost function f for our instantaneous control problem satsifies the _optimality equation_ (or Bellman equation)

(1.2) $0 = [\Gamma f(x) - \gamma f(x) + h(x)] \wedge [f'(x)+r] \wedge [f'(x)-\ell]$.

where

$$\Gamma = \frac{1}{2} \sigma^2 \frac{\partial^2}{\partial x^2} + \mu \frac{\partial}{\partial x} \, ,$$

is the infinitesimal generator of the Brownian motion X. Note that (1.2) imposes three differential inequalities on f, plus the requirement that at least one of the inequalities be tight in each

state x. For those familiar with standard stochastic control theory,
(1.2) has a strange appearance, so we shall begin in §2 with a
heuristic derivation of this optimality equation. A precise
mathematical formation of the instantaneous control problem is then
given in §3. Using the ubiquitous change of variable formula (or
generalized Ito Formula) for semimartingales, we show in §4 that any
smooth solution f of (1.2) satisfies $f \leq k$ for all cost functions
k associated with feasible policies. If one can find a policy whose
cost function satisfies the optimality equation, it of course follows
that this policy is optimal. The cost function for a general control
limit policy is computed in §5, and then in §6 we show how to choose
the control limits so that our optimality equation is satisfied by the
associated cost function. Finally, §7 discusses applications of the
instantaneous control problem and some correlative references.

2. Heuristic Derivation of the Optimality Equation

To simplify discussion, we assume in this section that $h(x) < \infty$
for all $x \in R$ (the real line), and the letters α and β will be
used here with new (temporary) meanings. Suppose that, in the
stochastic control problem described in §1, the controller can only
push right or left at a _rate_ which is not to exceed $\theta < \infty$. We then
have a more or less standard stochastic control problem, which can be
stated in the following unconventional form. A policy is a pair of
processes (R,L) having the form

(2.1) $R_t = \int_0^t \alpha_s \, ds$ and $L_t = \int_0^t \beta_s \, ds$, $t \geq 0$,

where α and β are non-anticipating with respect to X and satisfy $0 \leq \alpha, \beta \leq \theta$. The content process under policy (R,L) is $Z_t = X_t + R_t - L_t$, and we define the associated cost function $k(x)$ as in §1. Let $f(x)$ be the pointwise infimum of all such cost functions (the mimimal cost function). Under mild assumptions on h it can be shown that f is twice continuously differentiable and satisfies the optimality equation

(2.2) $0 = \inf\{\Gamma f(x) - \gamma f(x) + (\alpha - \beta) f'(x) + r\alpha + \ell\beta + h(x)\}$

          $= \inf\{(\Gamma f - \gamma f + h)(x) + \alpha[r + f'(x)] + \beta[\ell - f'(x)]\}$,

where the infimum is taken over all real numbers $\alpha, \beta \in [0,\theta]$. Problems of this type, where the controller has the ability to alter the drift of a diffusion process at some cost, have been treated by Mandl [9], Krylov [8], Fleming and Rishel [4], Gihman and Skorohod [5], and a number of others.

Since $f \in \mathscr{C}^2(\mathbb{R})$, the infimum in (2.2) is attained, and the minimizing values for α and β are

          $\alpha^*(x) = \theta 1_A(x)$ where $A = \{x \in R: \ r + f'(x) \leq 0\}$,

          $\beta^*(x) = \theta 1_B(x)$ where $B = \{x \in R: \ \ell - f'(x) \leq 0\}$.

6

This is a <u>bang-bang policy</u>. In each state x, each available control mode (pushing right and pushing left) is employed at either the maximum possible intensity θ or else at the minimum intensity of zero. The content process Z^* associated with the optimal policy satisfies the stochastic differential equation

$$(2.3) \qquad Z_t^* = X_t + \int_0^t \alpha^*(Z_t^*)ds - \int_0^t \beta^*(Z_s^*)ds$$

$$= X_t + \theta \int_0^t 1_A(Z_s^*)ds - \theta \int_0^t 1_B(Z_s^*)ds \quad ,$$

the optimal policy (R^*, L^*) being given by the last two terms on the right side of (2.3). Nothing said so far depends on any particular structure of h. Assuming that h is convex with $h(x) \to \infty$ as $|x| \to \infty$, it can be shown that f is convex itself, with $f(x) \to \infty$ as $|x| \to \infty$, so that $A = (-\infty, a]$ and $B = [b, \infty)$ for some parameters a and b $(-\infty < a < b < \infty)$.

Letting $\theta \to \infty$ in an attempt to approximate the instantaneous control problem of §1, this suggests the optimality of a control lir'' policy. Starting from any state x, we should either apply no control at all initially, or push right at the maximum possible rate, or else push left at the maximum rate. In the limiting problem, the latter two actions amount to instantaneous (jump) displacement, either right or left. With this motivation, we now use an infinitesimal argument to derive the optimality equation with instantaneous control, considering infinitesimal elements of space rather than time.

7

Let f be the minimal cost function for the instantaneous
control problem, fix a starting state x , and consider a small
surrounding interval $[x-\varepsilon, x+\varepsilon]$. The preceeding discussion suggests
that we should either jump immediately to $x+\varepsilon$ and proceed optimally
from there, jump immediately to $x-\varepsilon$ and proceed optimally from
there,or else apply no control up to time

$$T(\varepsilon) \equiv \inf\{\ t \geq 0: \ |X_t - X_0| = \varepsilon\}\ ,$$

and proceed optimally thereafter. Under the first option, our total
expected discounted cost is

(2.4) $$r\varepsilon + f(x+\varepsilon) = f(x) + [r + f'(x)]\varepsilon + o(\varepsilon)\ ,$$

under the second it is

(2.5) $$\ell\varepsilon + f(x-\varepsilon) = f(x) + [\ell - f'(x)]\varepsilon + o(\varepsilon)\ ,$$

and under the last it is

(2.6) $$E_x[\int_0^{T(\varepsilon)} e^{-\gamma t} h(X_t)dt + e^{-\gamma T(\varepsilon)} f(X_{T(\varepsilon)})]$$

$$= f(x) + [\Gamma f(x) - \gamma f(x) + h(x)]\ E_x[T(\varepsilon)] + o(E_x[T(\varepsilon)])$$

$$\approx f(x) + [\Gamma f(x) - \gamma f(x) + h(x)]\ (\varepsilon/\sigma)^2 + o(\varepsilon^2)\ \ .$$

8

In writing (2.6), we have used the fact that

$$E_x[f(X_{T(\epsilon)}) - f(x)]/E_x[T(\epsilon)] \to \Gamma f(x) \quad \text{as} \quad \epsilon \downarrow 0 ,$$

and that $\sigma^2 E_x[T(\epsilon)]/\epsilon^2 \to 1$ as $\epsilon \downarrow 0$. Now to <u>minimize</u> our expected discounted cost starting from x, we want to take the smallest of (2.4) - (2.6), meaning that

$$(2.7) \qquad f(x) = \min\{f(x) + [\Gamma f(x) - \gamma f(x) + h(x)] (\epsilon/\sigma)^2 + o(\epsilon^2) ,$$

$$f(x) + [r + f'(x)]\epsilon \qquad\qquad + o(\epsilon) ,$$

$$f(x) + [\ell - f'(x)]\epsilon \qquad\qquad + o(\epsilon)\}.$$

Substracting $f(x)$ from both sides of (2.7), and letting $\epsilon \downarrow 0$, we conclude that

$$0 = [\Gamma f(x) - \gamma f(x) + h(x)] \wedge [r + f'(x)] \wedge [\ell - f'(x)] ,$$

which is precisely the optimality equation (1.2).

3. Problem Formulation

The data for our problem are a drift parameter μ, a variance parameter $\sigma^2 > 0$, control cost parameters r and ℓ, an interest rate $\gamma > 0$, a compact state space $S \equiv [\alpha, \beta]$ and a convex holding cost function $h: S \to \mathbb{R}$. We assume $r + \ell > 0$, for otherwise the control problem would make no sense.

Let Ω be the space of all continuous functions $\omega: [0, \infty) \to \mathbb{R}$, which is usually denoted $C[0, \infty)$. Let $X_t: \Omega \to \mathbb{R}$ be the coordinate projection mapping $X_t(\omega) = \omega(t)$, $t \geq 0$ and $\omega \in \Omega$. Then $X = (X_t, t \geq 0)$ is simply the identity map $\Omega \to \Omega$. Let $\mathscr{F} = \sigma(X_t, t \geq 0)$ denote the smallest σ-field on Ω such that X_t is $\mathscr{F}$-measurable for each $t \geq 0$, and similarly let $\mathscr{F}_t = \sigma(X_s, 0 \leq s \leq t)$. Finally, for each $x \in S$, let P_x be the unique probability measure on $(\Omega, \mathscr{F})$ such that X is a Brownian motion with drift μ, variance σ^2 and starting state x under P_x, and let E_x be the associated expectation operator. A policy is defined as a pair of processes $R = (R_t, t \geq 0)$ and $L = (L_t, t \geq 0)$ such that

(3.1) $R(\omega)$ and $L(\omega)$ are right-continuous, non-negative

and non-decreasing for all $\omega \in \Omega$, and

(3.2) R_t and L_t and $\mathscr{F}_t$-measurable for all $t \geq 0$.

As usual, we summarize (3.2) by saying that R and L are _adapted_ to $(\mathscr{F}_t)$. We associate with policy (R, L) the _controlled process_ $Z = X + R - L$, and we say that (R, L) is a _feasible_ policy if

(3.3) $P_x(Z_t \in S$ for all $t \geq 0) = 1$ for all $x \in S$,

(3.4) $E_x[\int_0^\infty e^{-\gamma t} dR_t] < \infty$ for all $x \in S$,

and

(3.5) $E_x[\int_0^\infty e^{-\gamma t} dL_t] < \infty$ for all $x \in S$.

The integrals in (3.4) and (3.5) are defined for each fixed ω in the ordinary Lebesgue-Stieltjes sense over $[0,\infty)$, with the convention that $R_t = L_t = 0$ for $t < 0$. Thus the integral in (3.4), for example, equals the sum of R_0 and an integral over $(0,\infty)$. This same notational convention will be used later without comment. We associate with a feasible policy (R,L) the <u>cost function</u>

$$k(x) = E_x\{\int_0^\infty e^{-\gamma t}[h(Z_t)dt + rdR_t + \ell dL_t]\} , \quad x \in S ,$$

and (R,L) is said to be <u>optimal</u> if $k(x)$ is minimal (among the cost functions for feasible policies) for each $x \in S$. By defining feasibility via (3.3)-(3.5) we are implicitly setting $h = \infty$ outside S and then restricting attention to policies that have finite expected discounted cost for every starting state x. We could still enrich our definitional system to include all possible starting states $x \in \mathbf{R}$, but for x lying outside S, the feasibility restriction (3.5) would obviously require that either R or L have a jump at

11

$t = 0$ so as to ensure $Z_0 \in S$. By restricting attention to starting states $x \in S$ we avoid some irritating complications without any significant loss of generality.

This is the most concrete possible formulation of the decision problem described informally in §1. By taking $\Omega = C[0,\infty)$ and $X(\omega) = \omega$, we formally express the fact that our decision-maker observes nothing of relevance other than the sample path of X, and (3.2) expresses the requirement that his actions over the time interval $[0,t]$ depend only on the observed values of X_s, $0 \le s \le t$.

4. An Application of the Generalized Ito Formula

Until further notice, let (R,L) be a fixed feasible policy and $x \in S$ a fixed initial state. In the usual way, we denote by $\Delta R_t = R_t - R_{t-}$ the jump of R at time t, recalling that a right-continuous function with finite left limits can have only countably many points of discontinuity. As in §3, we take $R_t = 0$ for $t < 0$, so that $\Delta R_0 = R_0$. The same convention is used in extending the definition of ΔL_t to $t = 0$. It will be convenient to denote by ρ and λ the continuous parts of R and L respectively, meaning that

$$(4.1) \qquad \rho_t \equiv R_t - \sum_{0 \le s \le t} \Delta R_s \qquad \text{and} \qquad \lambda_t \equiv L_t - \sum_{0 \le s \le t} \Delta L_s \, ,$$

12

for $t \geq 0$. Obviously ρ and λ are continuous and non-decreasing with $\rho_0 = \lambda_0 = 0$. As before, let $Z = X+R-L$.

Now fix $f \in \mathscr{C}^2(S)$ and denote by $f(Z)$ the process $(f(Z_t),$ $t \geq 0)$. Then $\Delta f(Z)_t = f(Z_t) - f(Z_{t-})$ and we extend this to $t = 0$ with the useful convention

$$\Delta f(Z)_0 \equiv f(Z_0) - f(X_0) \quad .$$

Finally, let $\Gamma = 1/2 \, \sigma^2 \, \partial^2/\partial x^2 + \mu \, \partial/\partial x$ as in §1, and let $T > 0$ be fixed.

(4.2) **Proposition.** With the assumptions and definitions above,

$$E_x[e^{-\gamma T} f(Z_T)] = f(x) + E_x\left[\int_0^T e^{-\gamma T}(\Gamma f - \gamma f)(Z_t)dt\right]$$

$$+ E_x\left[\int_0^T e^{-\gamma t} f'(Z_t) \, d(\rho-\lambda)_t\right]$$

$$+ E_x\left[\sum_{0 \leq t \leq T} e^{-\gamma t} \Delta f(Z)_t\right]$$

Remark. From (3.3) - (3.5) and the fact that $f \in \mathscr{C}^2(S)$ it follows that all of the expectations appearing in (4.2) exist and are finite.

Proof. This is a direct application of the change of variable formula for semimartingales, but to make connection with the literature on that subject we must enrich our set-up slightly. Remembering that

13

$x \in S$ has been fixed, let $(\Omega, \mathscr{F}^*, P_x)$ be the completion of $(\Omega, \mathscr{F}, P_x)$, cf. Williams [11, p. 16], and for each $t \geq 0$ let $\mathscr{F}^*_t$ be formed from $\mathscr{F}_t$ by adding to it all $\Lambda \in \mathscr{F}^*$ such that $P_x(\Lambda) = 0$. It is well known that the filtration $(\mathscr{F}^*_t, t \geq 0)$ is right continuous, and hence the filtered probability space $(\Omega, \mathscr{F}^*, P_x, (\mathscr{F}^*_t, t \geq 0))$ satisfies the _usual conditions_ imposed by Meyer [10] in his treatment of stochastic integration and the change of variable formula. Our processes X, L, R and Z are all adapted to $(\mathscr{F}_t)$ and thus also to $(\mathscr{F}^*_t)$. When we use the terms adapted, martingale, stopping time, etc., later in this proof, the underlying filtration is understood to be $(\mathscr{F}^*_t)$.

It will be convenient to represent X in the form $X_t = X_0 + \sigma W_t + \mu t$ where W is a standard Wiener process starting at zero. Then

$$(4.3) \qquad Z_t = \sigma W_t + V_t \qquad \text{where} \qquad V_t \equiv X_0 + R_t - L_t + \mu t .$$

From the definitive properties of R and L we see that V is a VF (finite variation) process, and W is of course a martingale, so Z is a semimartingle. (We follow Meyer [10] in all of our terminology concerning martingales and related theory.) Then the change of variable formula (or generalized Ito formula) gives us

$$(4.4) \qquad f(Z_T) = f(Z_0) + \int_0^T f'(Z_{t-})dZ_t + \frac{1}{2}\sigma^2 \int_0^T f''(Z_{t-})dt$$
$$+ \sum_{0 < t \leq T} [\Delta f(Z)_t - f'(Z_{t-})\Delta Z_t] ,$$
$$\equiv f(Z_0) + I_1(T) + I_2(T) + \Sigma(T) ,$$

cf. Meyer [10, p. 301]. Here $I_1(T)$ is a stochastic integral over $(0,T]$, and we have simplified the general form of $I_2(T)$ by using the fact that σW is the so-called continuous martingale part of Z and $\langle \sigma W, \sigma W \rangle_t = \sigma^2 t$. Using (4.1) and (4.3), we have

$$(4.5) \qquad I_1(T) = \int_0^T f'(Z_{t-}) \, (\sigma dW_t + d\rho_t - d\lambda_t + \mu dt)$$

$$+ \sum_{0 < t \leq T} f'(Z_{t-}) \Delta Z_t \; .$$

Now we can replace $f'(Z_{t-})$ by $f'(Z_t)$ in the integral on the right side of (4.5), because the integrator is continuous, and a similar statement holds for $I_2(T)$. Thus, substituting (4.5) into (4.4) and combining similar terms, we have

$$(4.6) \qquad f(Z_T) = f(Z_0) + \sigma \int_0^T f'(Z_t) \, dW_t + \int_0^T f'(Z_t) \, d(\rho-\lambda)_t$$

$$+ \int_0^T \Gamma f(Z_t) dt + \sum_{0 < t \leq T} \Delta f(Z_t) \; .$$

Now let $Y_t = \exp(-\gamma t)$, $t \geq 0$. Because Y is a continuous VF process, the general integration by parts formula stated on page 303 of Meyer [10] simplifies to give us (in this equation, square brackets denote quadratic variation)

$$(4.7) \quad Y_T \, f(Z_T) = Y_0 f(Z_0) + \int_0^T Y_{t-} df(Z)_t + \int_0^T f(Z_{t-}) dY_t + [Y, f(Z)]_T$$

$$= Y_0 f(Z_0) + \int_0^T Y_t df(Z)_t + \int_0^T f(Z_t) \, dY_t \; ,$$

which is equivalent to

$$(4.8) \qquad e^{-\gamma T} f(Z_T) = f(Z_0) + \int_0^T e^{-\gamma t} df(Z)_t - \gamma \int_0^T e^{-\gamma t} f(Z_t) dt \ .$$

Now we calculate $df(Z)_t$ from (4.4), substitute this into (4.8) and collect similar terms to get

$$(4.9) \quad e^{-\gamma T} f(Z_T) = f(Z_0) + \sigma \int_0^T e^{-\gamma t} f'(Z_t) \, dW_t$$

$$+ \int_0^T e^{-\gamma t} f'(Z_t) \, d(\rho - \lambda)_t + \int_0^T e^{-\gamma t} (\Gamma f - \gamma f)(Z_t) dt$$

$$+ \sum_{0 < t \leq T} e^{-\gamma t} \Delta f(Z)_t \ .$$

Next, because $\Delta f(Z)_0 \equiv f(Z_0) - f(X_0)$, we have

$$(4.10) \qquad f(Z_0) + \sum_{0 < t \leq T} e^{-\gamma t} \Delta f(Z)_t = f(X_0) + \sum_{0 \leq t \leq T} e^{-\gamma t} \Delta f(Z)_t \ .$$

We now substitute (4.10) into (4.9) and take E_x of both sides, observing that the Ito integral involving dW_t has zero expectation because its integrand is bounded. This yields equation (4.2) and thus completes the proof.

Maintaining the set-up for (4.2), we now define the cumulative discounted cost process associated with policy (R,L). Let

16

(4.11) $K_t \equiv \int_0^t e^{-\gamma s}[h(Z_s) + rdR_s + \ell dL_s]$, $t \geq 0$.

the second and third integrals being defined in the Lebesgue-Stieltjes
sense over $[0,t]$ with the usual convention at zero ($\Delta R_0 = R_0$ and
$\Delta L_0 = L_0$).

(4.12) <u>Corollary</u>. With the assumptions and definitions above,

$$E_x[K_T + e^{-\gamma T} f(Z_T)]$$

$$= f(x) + E_x\{\int_0^T e^{-\gamma t}(\Gamma f - \gamma f + h)(Z_t)dt\}$$

$$+ E_x\{\int_0^T e^{-\gamma t} [r + f'(X_t)] d\rho_t\}$$

$$+ E_x\{\int_0^T e^{-\gamma t}[\ell - f'(Z_t)]d\lambda_t\}$$

$$+ E_x\{\sum_{0 \leq t \leq T} e^{-\gamma t}[\Delta f(Z)_t + r\Delta R_t + \ell\Delta L_t]\} \quad .$$

(4.13) <u>Remark</u>. For future reference, we express the right side of
(4.12) as $f(x) + E_x[I_1(T) + I_2(T) + I_3(T) + \Gamma(T)]$.

<u>Proof.</u> This follows immediately from (4.2) and (4.11), using the
identities $dR = d\rho + \Delta R$ and $dL = d\lambda + \Delta L$ in (4.11).

For an interpretation of (4.12), imagine that you have responsibility for operating the storage system described in §1 and have tentatively decided to use the policy (R,L). Further suppose that another person offers to relieve you of this responsibility under either of the following two arrangements.

(a) You may pay $f(x)$ dollars at time zero and avoid all future control and holding costs.

(b) You may employ policy (R,L) up to time T, absorbing the control and holding costs incurred during that period, then make a payment of $f(Z_T)$ at time T and be relieved of all control responsibilities thereafter.

Corollary (4.12) gives an expression for your expected discounted cost under plan (b). Now suppose that f satisfies

$$(4.14) \qquad \Gamma f - \gamma f + h \geq 0 \qquad \text{on } S ,$$

$$(4.15) \qquad r + f' \geq 0 \qquad \text{on } S ,$$

and

$$(4.16) \qquad \ell - f' \geq 0 \qquad \text{on } S .$$

Using the notational convention (4.13), it is clear that (4.14) implies $E_x[I_1(T)] \geq 0$, (4.15) implies $E_x[I_2(T)] \geq 0$, and (4.16) implies $E_x[I_3(T)] \geq 0$. Furthermore, (4.15) and (4.16) together imply $E_x[\Sigma(T)] \geq 0$ as follows. Suppose $\Delta R_t > 0$ and $\Delta L_t = 0$. Then $\Delta Z_t = \Delta R_t$ and we have

$$\Delta f(Z)_t + r\Delta R_t + \ell\Delta L_t = f(Z_t) - f(Z_t - \Delta R_t) + r\Delta R_t$$

$$= \int_{Z_t - \Delta R_t}^{Z_t} [f'(y) + r]\,dy \geq 0 , \quad \text{by (4.15)} .$$

From (4.16) we get a similar inequality for times t where $\Delta R_t = 0$ and $\Delta L_t > 0$. Finally, (4.15) and (4.16) together imply a similar inequality for times t with $\Delta R_t > 0$ and $\Delta L_t > 0$, because (and only because) we have assumed $r + \ell > 0$. So we find that (4.14) – (4.16) imply

$$(4.17) \qquad\qquad E_x[K_T + e^{-\gamma T} f(Z_T)] \geq f(x) ,$$

which means that plan (b) above is inferior to plan (a) for any choice of T (and regardless of the starting state x). Letting $T \to \infty$ in (4.17) gives $k(x) \geq f(x)$, since f is bounded on S. Since (R,L) and x were arbitrary, we then have the following.

(4.18) <u>Corollary.</u> If $f \in \mathscr{C}^2(S)$ satsifies (4.14) – (4.16), then $f \leq k$ for any cost function k associated with a feasible policy.

Corollary (4.18) is the only result from this section that will be used later, but we should say at least a few words to connect our basic identity (4.12) with the optimality equation (1.1) and the

general notion of policy improvement. Suppose that f is the cost

function for a feasible policy (R*,L*) that we want to test for

optimality. The left side of (4.12) gives the expected discounted

cost when we use an alternate policy (R,L) up to time T and employ

(R*,L*) thereafter, with T playing the role of time zero and Z_T

viewed as the initial state of the control problem. (To make this

last phrase precise, one must introduce shift operators.) For (R*,

L*) to be an optimal policy, it is necessary and sufficient that all

such attempts to improve (R*,L*) through hybridization fail, meaning

that (4.17) holds for every x, every stopping time T, and every

feasible policy (R,L). Combining this with (4.12), it can be shown

that (4.14)-(4.16) are <u>necessary and sufficient</u> for the optimality of

(R*,L*). Finally, from (4.12) and the fact that f is (by

assumption) the cost function for a feasible policy, it can be shown

that at least one of the inequalities (4.14) - (4.16) is tight at each

point x ϵ S, meaning that if satisfies the optimality equation

$$(4.19) \qquad 0 = [(\Gamma f - \gamma f + h) \wedge (r + f') \wedge (\ell - f')](x) , \qquad x \epsilon S ,$$

which appeared earlier as (1.1). To repeat, if f ϵ $\mathscr{C}^2(S)$ is the

cost function for a feasible policy, then (4.19) is necessary and

sufficient for the optimality of that policy, but only the sufficiency

has been proved rigorously.

5. Control Limit Policies

Let a and b be fixed throughout this section, with $\alpha \le a < b \le \beta$. We want to construct the policy (R,L) that enforces these control limits, and then calculate the associated cost function. These are essentially known results for one-dimensional Brownian motion with reflecting barriers, but we do not know of any textbook treatment that presents them in a form suitable for our purposes.

(5.1) **Proposition.** For each $\omega \in \Omega$ there exists a unique pair of functions $R(\omega) = \{R_t(\omega), t \ge 0\}$ and $L(\omega) = \{L_t(\omega), t \ge 0\}$ which jointly satisfy

(5.2)
$$R_t(\omega) = \sup_{0 \le s \le t} [a - X_s(\omega) + L_s(\omega)]^+ , \qquad t \ge 0 ,$$

(5.3)
$$L_t(\omega) = \sup_{0 \le s \le t} [X_s(\omega) + R_s(\omega) - b]^+ , \qquad t \ge 0 .$$

Both $R(\omega)$ and $L(\omega)$ are continuous and non-decreasing, with $R_0(\omega) = [a - X_0(\omega)]^+$ and $L_0(\omega) = [X_0(\omega) - b]^+$.

Proof. Let us first prove the last statement, taking R and L to be any two functions which jointly satisfy (5.2) - (5.3). (The dependence on ω will be suppressed throughout this proof.) If R_0 and L_0 were both positive, then we would have $R_0 = a - X_0 + L_0$ by (5.2) and $L_0 = X_0 + R_0 - b$ by (5.3), which implies $a = b$, a contradiction. In exactly the same way, if $\Delta R_t > 0$ and $\Delta L_t > 0$ for some

21

$t > 0$, then the suprema in (5.2) and (5.3) would both be achieved at $s = t$, implying $R_t = a - X_t + L_t$ and $L_t = X_t + R_t - b$, and again we arrive at the contradiction $a = b$. So R and L cannot jump simultaneously, and then (5.2)-(5.3) and the continuity of X imply that they have no jumps at all.

We now construct a solution of (5.2)-(5.3) by successive approximations. Beginning with the trial solution $R_t^0 = L_t^0 = 0$, $t \geq 0$, let

$$(5.4) \qquad R_t^{n+1} = \sup_{0 \leq s \leq t} [a - X_s + L_s^n]^+ ,$$

and

$$(5.5) \qquad L_t^{n+1} = \sup_{0 \leq s \leq t} [X_s + R_s^n - b]^+ ,$$

for $n = 0, 1, \ldots$ and $t \geq 0$. Observe that $R_t^1 \geq R_t^0$ and $L_t^1 \geq L_t^0$, and hence (by induction) that R_t^n and L_t^n are increasing in n for each fixed t. So we have

$$(5.6) \qquad R_t^n \uparrow R_t \qquad \text{and} \qquad L_t^n \uparrow L_t \qquad \text{as} \qquad n \uparrow \infty, \; t \geq 0 ,$$

and one can easily verify that the convergence in (5.6) is obtained in a finite number of iterations for each fixed t. Thus R and L are finite valued and jointly satisfy (5.2)-(5.3).

For uniqueness, let (R', L') be another (distinct) solution of (5.2)-(5.3). We have already seen that R' and L' must both be

22

continuous with $R_0' = R_0$ and $L_0' = L_0$, and it's obvious from the construction above that $R' \geq R$ and $L' \geq L$. Let $T > 0$ be the infinium of those $t > 0$ at which either $R_t' > R_t$ or $L_t' > L_t$. By continuity, we have $R_T' = R_T$ and $L_T' = L_T$, and either R' or L' must increase at T. If T were a point of increase for both, then (5.2) and (5.3) would give us $R_T = a - X_T + L_T$ and $L_T = X_T + R_T - b$ respectively, which yields the contradiction $a = b$. So we conclude that exactly one of the pair (R', L') increases at T. Suppose it is R', implying that $L' = L$ over $[0, T+\varepsilon]$ for sufficiently small ε. But then $R' = R$ over $[0, T+\varepsilon]$ by (5.2), which contradicts the definition of T. In the same way, we cannot have R' flat and L' increasing at T, so the proof of uniqueness is complete.

(5.7) **Proposition.** Let $R(\omega)$ and $L(\omega)$ be as in (5.1), and set $Z = X + R - L$. The processes R, L and Z are adapted and satisfy

$$(5.8) \qquad a \leq Z_t \leq b , \qquad\qquad t \geq 0 ,$$

$$(5.9) \qquad \int_0^t (Z_s - a) dR_s = 0 , \qquad\qquad t \geq 0 ,$$

$$(5.10) \qquad \int_0^t (b - Z_s) dL_s = 0 , \qquad\qquad t \geq 0 .$$

Remark. One may paraphrase (5.9) by saying that R increases only when $Z = a$. With our usual convention, (5.9) yields $(Z_0 - a)R_0 = 0$ when specialized to $t = 0$. Similar statements hold for (5.10).

Proof. The adaptedness is immediate from our construction (5.4)-(5.6) of R and L, while (5.8)-(5.10) follow directly from (5.2)-(5.3).

It can further be shown that the _unique_ pair of functions (R,L) satisfying (5.8)-(5.10), with $Z \equiv X+R-L$, is that constructed in the proof of (5.1). This means that the characterizations of (R,L) given in Propositions (5.1) and (5.7) are completely equivalent. We observed earlier that the convergence (5.6) is obtained in a finite number of iterations for each fixed t, which makes it possible to write out a general (and very messy) recursive formula for R and L in terms of a sequence of stopping times $\{T_n\}$. This was done in [6] for the case $a = 0$, but the only relevant properties of the resulting pair (R,L) are those expressed in (5.7). It can be shown that R (respectively L) is the _local time_ of the diffusion process Z at the boundary a (respectively b), but we shall have no need for this fact.

(5.11) _Proposition._ Suppose that $k \in \mathscr{C}^1(S)$ is _twice_ continuously differentiable on [a,b] and satisfies

$$(5.12) \qquad \Gamma k(x) - \gamma k(x) + h(x) = 0 \ , \qquad a \leq x \leq b \ ,$$

$$(5.13) \qquad k'(x) + r = 0 \ , \qquad \alpha \leq x \leq a \ ,$$

24

(5.14) $\qquad\qquad k'(x) - \ell = 0$, $\qquad b \leq x \leq \beta$.

Then

$$k(x) = E_x \{ \int_0^\infty e^{-\gamma t} [h(Z_t) dt + r dR_t + \ell dL_t] \} , \qquad x \in S .$$

Remark. This of course shows that there is at most one k satisfying the stated hypotheses, and we shall exhibit a solution (or rather the solution) shortly.

Proof. First fix a starting state $x \in [a,b]$. Defining the cumulative discounted cost K_t as in (4.11), we need to prove that $E_x(K_\infty) = k(x)$. Fixing $T > 0$, we shall apply Corollary (4.12) with $k \in \mathscr{C}^2[a,b]$ replacing $f \in \mathscr{C}^2[\alpha,\beta]$. Since L and R have no jumps when $X_0 = x$, we have $\rho = R$ and $\lambda = L$, and (4.12) yields

(5.15) $\qquad E_x[K_T + e^{-\gamma T} k(Z_T)]$

$$= k(x) + E_x \{ \int_0^T e^{-\gamma t} (\Gamma k - \gamma k + h)(Z_t) dt \}$$

$$+ E_x \{ \int_0^T e^{-\gamma t} [r + k'(Z_t)] dR_t \}$$

$$+ E_x \{ \int_0^T e^{-\gamma t} [\ell - k'(Z_t)] dL_t \} .$$

Since $P_x\{a \leq Z_t \leq b$ for all $t \geq 0\} = 1$, the second term on the right side of (5.15) vanishes by (5.12). Next, (5.9) says that R increases only when $Z = a$, so the third term on the right side of (5.15) is

$$E_x\{\int_0^T e^{-\gamma t}[r+k'(a)]dR_t\} \quad ,$$

which vanishes by (5.13). Similarly, the final term on the right side of (5.15) vanishes by (5.10) and (5.14). Letting $T \to \infty$ in (5.15), and using the boundedness of $k(Z_T)$, we thus obtain $E_x(K_\infty) = k(x)$ as desired.

If $a = \alpha$ and $b = \beta$, there is nothing left to prove. Next suppose $b < \beta$ and consider a starting state $x \in (b,\beta]$. From the *construction of* (R,L) we have

(5.16)
$$E_x(K_\infty) = E_x(\ell L_0) + E_b(K_\infty)$$
$$= \ell(x-b) + E_b(K_\infty) \ .$$

The first part of the proof shows that $E_b(K_\infty) = k(b)$, and $\ell(x-b)+k(b) = k(x)$ by (5.14), so (5.16) reduces to $E_x(K_\infty) = k(x)$ as desired. A similar argument, using (5.13), gives $E_x(K_\infty) = k(x)$ for $\alpha \leq x < a$, which completes the proof.

We conclude the section by constructing a solution for the ordinary differential equation (5.12)-(5.14). To emphasize the

26

dependence of this solution on the control limits a and b, we denote it $k_{ab}(x)$. First, let $g \in \mathscr{C}^2(S)$ be the unique solution of

(5.7) $\Gamma g(x) - \gamma g(x) + h(x) = 0$, $\alpha \leq x \leq \beta$,

(5.8) $g(\alpha) = g(\beta) = 0$.

It is well known that exactly one such g exists, and it can be written explicitly as an integral involving a known Green's function. Next, setting

$$\theta_1 = [-\mu + (\mu^2 + 2\gamma\sigma^2)^{1/2}]/\sigma^2 \quad ,$$

$$\theta_2 = [-\mu + (\mu^2 + 2\gamma\sigma^2)^{1/2}]/\sigma^2 \quad ,$$

$$c_1 = \exp(\theta_2 a)/\theta_1 [\exp(\theta_1 b + \theta_2 a) - \exp(\theta_1 a + \theta_2 b)] \ ,$$

$$c_2 = \exp(\theta_1 a)/\theta_2 [\exp(\theta_1 a + \theta_2 b) - \exp(\theta_1 b + \theta_2 a)] \ ,$$

$$d_1 = \exp(\theta_2 b)/\theta_1 [\exp(\theta_1 a + \theta_2 b) - \exp(\theta_1 b + \theta_2 a)] \ ,$$

$$d_2 = \exp(\theta_1 b)/\theta_2 [\exp(\theta_1 b + \theta_2 a) - \exp(\theta_1 a + \theta_2 b)] \ ,$$

we define

$$\phi_1(x) = c_1 e^{\theta_1 x} + c_2 e^{\theta_2 x} \; , \qquad\qquad a \leq x \leq b \; ,$$

$$\phi_2(x) = d_1 e^{\theta_1 x} + d_2 e^{\theta_2 x} \; , \qquad\qquad a \leq x \leq b \; ,$$

$$\psi(x) = g(x) + [\ell - g'(b)] \phi_1(x) - [r + g'(a)] \phi_2(x) \; , \quad a \leq x \leq b \; ,$$

and finally

$$k_{ab}(x) = \begin{cases} \psi(a) + (a-x)r \; , & \alpha \leq x \leq a \; , \\[2mm] \psi(x) \; , & a \leq x \leq b \; , \\[2mm] \psi(b) + (x-b)\ell \; , & b \leq x \leq \beta \; . \end{cases}$$

One can easily verify that

$$\Gamma\phi_1(x) - \gamma\phi_1(x) = \Gamma\phi_2(x) - \gamma\phi_2(x) = 0 \; , \quad a \leq x \leq b \; ,$$

$$\phi_1'(a) = \phi_2'(b) = 0 \qquad \text{and} \qquad \phi_1'(b) = \phi_2'(a) = 1 \; .$$

From this and (5.7)-(5.8) it follows directly that k_{ab} is the desired solution of (5.12)-(5.14). Observe that k_{ab}' is continuous on all of $[\alpha, \beta]$ and that k_{ab}'' is continuous everywhere except possibly at the control limits a and b.

In this construction of k_{ab} one can start with <u>any</u> function $g \in \mathscr{C}^2(S)$ satisfying the main equation (5.7). For concreteness we have specified one particular solution via the boundary conditions (5.8). For future reference, note that our choice of g does not depend on the control limits a and b.

6. Optimal Control Limits

Maintaining the notation of the previous section, let $f = k_{ab}$ be the cost function for a control limit policy with $\alpha \leq a < b \leq \beta$. Recall that f is continuously differentiable on S and twice continuously differentiable except possibly at a and b. We would like to find conditions (on a and b) under which f satisfies the optimality conditions (4.14)-(4.16).

The first thing to note is that f can only satisfy (4.14)-(4.16) if it is twice continuously differentiable, meaning that $f''(x) \to 0$ as $x \downarrow a$ if $\alpha < a$, and $f''(x) \to 0$ as $x \uparrow b$ if $b < \beta$. To see this, suppose first that $\alpha < a$ and that $f''(a+) > 0$. Since $(\Gamma f - \gamma f + h)(a+) = 0$, and f, f' and h are continuous at a while $f''(a-) = 0$, we have $(\Gamma f - \gamma f + h)(a-) < 0$, which violates (4.14). Now suppose, on the other hand, that $f''(a+) < 0$. Since f' is continuous at a with $f'(a) = -r$, this implies $f'(a+\epsilon) < -r$ for $\epsilon > 0$ sufficiently small, which violates (4.15). So continuously of f'' at the lower control limit is <u>necessary</u> if f is to satisfy (4.14)-(4.16), and a similar analysis holds at the upper control limit. We turn now to the matter of sufficiency.

(6.1) <u>Proposition</u>. Suppose that $f = k_{ab}$ is twice continuously differentiable on S with $-r \leq f' \leq \ell$. Then f satisfies the optiality conditions (4.14)-(4.16).

<u>Proof</u>. If $a = \alpha$ and $b = \beta$, the conclusion is automatic. Suppose then that $a > \alpha$. Defining

$$\psi(x) \equiv \Gamma f(x) - \gamma f(x) + h(x)$$
$$= \frac{1}{2} \sigma^2 f''(x) + \mu f'(x) - \gamma f(x) + h(x) , \qquad x \epsilon S ,$$

we have $\psi(x) = 0$ for $a \le x \le b$ and need to prove that $\psi(x) \ge 0$ for $\alpha \le x \le a$. (The proof that $\psi(x) \ge 0$ for $b \le x \le \beta$ if $b < \beta$ is virtually identical, so we delete it.) Let us define

$$\phi(x) \equiv [h(x) - h(a)] + \gamma r(x-a) ,$$

$$\theta(x) \equiv [\Gamma f(x) - \Gamma f(a)] - \gamma[f(x) - f(a)] + \gamma f'(a) (x-a) .$$

Remembering that $\psi(a) = 0$, $\Gamma f(x) = \Gamma f(a)$ for $\alpha \le x \le a$, $f(x) = f(a) + r(a-x)$ for $\alpha \le x \le a$, and $f'(a) = -r$, we then have

$$(6.2) \qquad \psi(x) = \psi(x) - \psi(a) = \begin{cases} \phi(x) , & \text{if } \alpha \le x \le a \\ \\ \phi(x) + \theta(x) , & \text{if } a \le x \le \beta \end{cases}$$

Next, since $f \epsilon \mathscr{C}^2(S)$ by assumption, $f''(a) = 0$, $f'(a) = -r$ and $f'(x) \ge -r$ for all x by assumption, it must be that $f''(a+\xi) \ge 0$ for all $\xi \ge 0$ sufficiently small. From Taylor's theorem and the definition of $\theta(\cdot)$, we then have the following: for each $\delta > 0$ there exists an $\epsilon > 0$ such that

30

(6.3) $\theta(x) \geq -(x-a)\delta$,   for $a \leq x \leq a+\varepsilon$.

But $\phi(x) = 0$ for $a \leq x \leq b$, so (6.2) and (6.3) together imply

(6.4) $\phi(x) \leq (x-a)\delta$,   for $a \leq x \leq a+\varepsilon$.

Convexity of h implies convexity of ϕ, and obviously $\phi(a) = 0$, so (6.4) implies

(6.5) $\phi(x) \geq -(a-x)\delta$,   for $\alpha \leq x \leq a$

Since $\delta > 0$ was arbitrary, this gives $\phi(x) \geq 0$ for $\alpha \leq x \leq a$ and hence $\psi(x) \geq 0$ for $\alpha \leq x \leq a$ by (6.2). This completes the proof of the proposition.

We now construct a control limit policy whose cost function satisfies the hypotheses of Proposition (6.1). For each $a \in [\alpha, \beta)$ let

$$b^*(a) = \sup\{b \in (a,\beta]: k'_{ab}(x) \leq \ell, \ a \leq x \leq b\}.$$

From the explicit formula for k_{ab} given in §5, it follows easily that $b^*(a) > a$. Hereafter let $k_a \equiv k_{ab^*(a)}$ for $a \in [\alpha,\beta)$. Next define

$$a^* = \inf\{a \in [\alpha,\beta): k_a'(x) \geq -r,\ a \leq x \leq b^*(a)\} \quad .$$

Again it follows easily from the formulas of §5 that $a^* < \beta$.
Hereafter we set $b^* \equiv b^*(a^*)$.

(6.6) **Proposition.** The cost function $f = k_{a^*b^*}$ satisfies the
hypotheses of (6.1), and thus the control limit policy with parameters
a^* and b^* is optimal.

Proof. The inequality $-r \leq f' \leq \ell$ is immediate from our construc-
tion. It remains to show that f'' is continuous, which means simply
that

(6.7) $\qquad\qquad f''(x) \to 0 \qquad$ as $\qquad x \downarrow a^* \qquad$ if $\qquad a^* > \alpha$,

(6.8) $\qquad\qquad f''(x) \to 0 \qquad$ as $\qquad x \uparrow b^* \qquad$ if $\qquad b^* < \beta$.

From the formulas of §5 it is immediate that k_{ab} and its first two
derivatives vary continuously with the parameters a and b, and it
is this continuity plus the definitions of a^* and b^* that one uses
in verifying (6.7)-(6.8). The verification is straightforward but
tedious, so we leave it as an exercise. Propositions (6.1) and (4.18)
give $f \leq k$ for any cost function k associated with a feasible
policy, which completes the proof of the proposition.

32

The approach that we have taken to determining an optimal policy does not work directly with the optimality equation (1.2), but the return function for our optimal control limit policy does in fact satisfy this relationship. Defining $f = k_{a*b*}$ as in (6.6), we have seen that f is the minimal cost function and that it satisfies each of the inequalities (4.14)-(4.16). Furthermore, from the definition and construction of k_{ab} given in §5, we see that (4.14) holds with equality on $[a*,b*]$, (4.15) holds with equality on $[\alpha,a*]$ and (4.16) holds with equality on $[b*,\beta]$. Thus,

(6.9) $0 = [\Gamma f(x) - \gamma f(x) + h(x)] \wedge [r + f'(x)] \wedge [\ell - f'(x)], \quad x \in S$,

as claimed. Adding the boundary conditions

(6.10) $r + f'(\alpha) = 0 \quad$ and $\quad \ell - f'(\beta) = 0$,

we believe that the <u>unique</u> function $f \in \mathscr{C}^2(S)$ satisfying (6.9)-(6.10) is the minimum cost function, but we have not attempted to prove this. The boundary conditions (6.10) are essential, incidentally, since there may exist $f \in \mathscr{C}^2(S)$ satisfying $\Gamma f - \gamma f + h = 0$ on S with $r + f'(x) > 0$ for all $x \in S$ or $\ell - f'(x) > 0$ for all $x \in S$. Such an f satisfies (6.9) but lies strictly below the minimum cost function.

33

7. Concluding Remarks

Two potential areas of application for our instantaneous control problem are cash management and production control. See Harrison and Taylor [6] for a discussion of these applications, oriented toward the specific holding cost function (1.1), and further references. For cash management problems, the non-linear (but convex) holding cost function

$$(7.1) \qquad h(x) = \begin{cases} hx , & \text{if } x \geq 0 \quad (h > 0) \\ \\ q|x| , & \text{if } x \leq 0 \quad (q > 0) \end{cases}$$

is also of practical importance. This arises when the firm can maintain a negative cash balance through short-term borrowing, and a similar holding cost structure occurs in production control problems where demand can be backlogged at some penalty cost. Motivated by the stochastic cash management problem, Constantinides and Richard [3] have studied the optimal control of Brownian Motion when the holding cost function is (7.1) and there are both fixed and proportional costs of control. This gives a problem of optimal impulse control [2], and they show the existence of an optimal policy characterized by four critical numbers. The applications of instantaneous control in production and inventory theory will be further developed in [7], using the results of this paper.

Two problems of instantaneous control, closely related to ours but much more difficult, have been solved in a beautiful recent paper

34

by Benes, Shepp and Witsemhausen [1]. Using our notational system,
one of their problems can be stated as follows: Find a pair of
controls (R,L) to minimize

$$E\left[\int_0^\infty e^{-\alpha t}(X_t + R_t - L_t)^2 \, dt\right] ,$$

subject to the constraint that $R_\infty + L_\infty \leq y < \infty$. Here one has no
explicit cost of control, but there is a finite limit on the total
amount of control that can be exerted over the infinite planning
horizon. The authors take essentially the same approach employed
here, using martingale methods (the generalized Ito formula) to verify
optimality of a candidate policy arrived at from certain heuristic
considerations. Their optimal policy has a much more complex form
that ours, however, so the argument is much more intricate. This is
the only previous paper we know of, other than [6], which explicitly
considers an instantaneous control problem for Brownian motion, as
opposed to control at a bounded rate or optimal impulse control.

References

[1] V.E. Benes, L.A. Shepp and H.S. Witsenhausern (1980), *Some Solvable Stochastic Control Problems*, Stochastics, Vol. 4, 39-83.

[2] A. Bensoussan and J.L. Lions (1975), Nouvelles Methods en Controle Impulsionnel, Appl. Math. Opt., Vol. 1, 289-312.

[3] G.M. Constantinides and S.F. Richard (1978), Existence of Optimal Simple Policies for Discounted-Cost Inventory and Cash Management in Continuous Time, Ops. Resch., Vol. 26, 620-636.

[4] W.H. Fleming and R.W. Rishel (1975), Deterministic and Stochastic Optimal Control, Springer-Verlag, Berlin.

[5] I.I. Gihman and A.V. Skorohod (1979), Controlled Stochastic Process, Springer-Verlag, Berlin.

[6] J.M. Harrison and A.J. Taylor (1978), Optimal Control of a Brownian Storage System, Stoch. Proc. Appl., Vol. 6, 179-194.

[7] J.M. Harrison (1981), A Production Control Problem and its Diffusion Limit, forthcoming.

[8] N.V. Krylov (1977), Controlled Processes of Diffusion Type (in Russian), Nauka, Moskow.

[9] P. Mandl (1968), Analytic Treatment of One-Dimensional Markov Processes, Springer-Verlag, Berlin.

[10] P.-A. Meyer (1976), Un Cours sur les Integrales Stochastiques, in Sem. de Prob. X, Lecture Notes in Mathematics 511, Springer-Verlag, Berlin.

[11] D. Williams (1979), Diffusions, Markov Processes and Martingales, Wiley, New York.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 199	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) INSTANTANEOUS CONTROL OF BROWNIAN MOTION		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) J. MICHAEL HARRISON and MICHAEL I. TAKSAR		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0561
9. PERFORMING ORGANIZATION NAME AND ADDRESS DEPT'S. OF OPERATIONS RESEARCH & STATISTICS STANFORD UNIVERSITY STANFORD, CALIFORNIA 94305		10. PROGRAM ELEMENT. PROJECT, TASK AREA & WORK UNIT NUMBERS NR-047-200
11. CONTROLLING OFFICE NAME AND ADDRESS OPERATIONS RESEARCH, CODE 434 OFFICE OF NAVAL RESEARCH ARLINGTON, VIRGINIA 22217		12. REPORT DATE OCTOBER 1981
		13. NUMBER OF PAGES 36
14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, If different from Report)

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

BROWNIAN MOTION,	STOCHASTIC CONTROL,
REFLECTING BARRIERS,	DIFFUSION PROCESSES,
INVENTORY AND PRODUCTION CONTROL,	STOCHASTIC CASH MANAGEMENT

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

PLEASE SEE OTHER SIDE

Instantaneous Control of Brownian Motion

J. Michael Harrison
Michael I. Taksar

Stanford University

Abstract

A controller continuously monitors a storage system, such as an inventory or bank account, whose content $Z = \{Z_t, t \geq 0\}$ fluctuates as a (μ, σ^2) Brownian motion in the absence of control. Holding costs are incurred continuously at rate $h(Z_t)$. At any time, the controller may instantaneously increase the content of the system, incurring a proporitional cost of r times the size of the increase, or decrease the content at a cost of ℓ times the size of the decrease. We consider the case where h is convex on a finite interval $[\alpha, \beta]$ and $h = \infty$ outside this interval. The objective is to minimize the expected discounted sum of holding costs and control costs over an infinite planning horizon.

It is shown that there exists an optimal control limit policy, characterized by two parameters a and b $(\alpha \leq a < b \leq \beta)$. Roughly speaking, this policy exerts the minimum amounts of control sufficient to keep $Z_t \in [a,b]$ for all $t \geq 0$. Put another way, the optimal control limit policy imposes on Z a lower reflecting barrier at a and an upper reflecting barrier at b. We do not give a full-blown algorithm for construction of the optimal control limits, but a computational scheme could easily be developed from our constructive proof of existence.

DATE
FILMED

2-8